DOCUMENT RESUME

ED 212 212 HE 014 692

AUTHOR Carlson, Robert V.
TITLE Are Evaluation Activities Worth the Effort?
PUB DATE 29 Oct 81
NOTE 38p.; Paper presented at the Annual Meeting of the American Society for Medical Technology (Rochester, NY, October 29, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Classification; *Educational Assessment; Evaluation Criteria; *Evaluation Methods; Formative Evaluation; Higher Education; Needs Assessment; *Program Evaluation; Research Methodology; *Standards; Summative Evaluation

ABSTRACT

The casual approach to evaluation, the interface of evaluation and research, and various approaches subsumed within evaluation are addressed, and the dimensions that determine the worth of engaging in evaluation activities are examined. Additionally, a statement about the state-of-the-art in regard to the technology of conducting evaluation studies is presented. The practice referred to as the "casual approach" appears to be a dominant style of a large percentage of organizational personnel. The following dimensions of a casual approach to evaluation are examined: face-to-face discussions, random observations, group discussions, and individual reflections. It is suggested that the potential strengths of the casual approach be strengthened with the tools and disciplines of evaluation and research. Comparisons are made between the methodologies of research, evaluation, evaluation research, and policy analysis/research in regard to purpose, process, products, and criteria for judging quality. Additionally, a taxonomy of eight specific evaluation approaches is presented. Factors that determine whether to conduct formal evaluation activities are identified, and utility standards, feasibility standards, proprietary standards, and accuracy standards for educational program evaluation are presented. In addition, steps of a program evaluation are detailed, and areas that need improved strategies are listed. A bibliography is appended. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Are Evaluation Activities Worth the Effort?

A Paper Presented
at the American Society for Medical Technology
Rochester, New York

October 29, 1981

Robert V. Carlson, Director
Center for Evaluation and Policy Research
College of Education and Social Services
University of Vermont
Burlington, Vermont

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Robert V. Carlson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

OUTLINE

1.0 Introduction - 1

2.0 As Compared to What? - - - - - - - - - - - - - - - - - 1

- Face-to-face discussions
- Random observation
- Group discussions
- Individual reflections

3.0 Evaluation Activities - - - - - - - - - - - - - - - - 3

4.0 Worth? - 4

- Need
- Resources
- Standards

5.0 A Further Note - 5

6.0 Summary - 6

Related Readings - 8

Exhibit 1 Interrelationship of Evaluation, Evaluation Research and Research to Program/Project Development

2 Comparison of Alternative Problem Solving Methodologies

3 A Taxonomy of Major Evaluation Approaches

4 Standards for Evaluations of Educational Programs

5 Standards for Program Evaluation

1.0 Introduction

Are evaluations worth the effort begs for a simple "yes" or "no" answer particularly from those who either must organize or carry out evaluations or those who must respond with time and effort to evaluation requirements. Unfortunately, a simple "yea" or "nay" hides the complexities that this question implies. It seems before we can resolve this issue, there are some related questions worthy of further exploration. These might be stated as follows:

1. What do we mean by evaluation activities?
2. How do we judge the worth or merit of anything?
3. And, finally, are evaluation activities worth the effort, as compared to what?

2.0 As Compared to What?

Taking the last question first, it may be well to examine what seems to be the most common alternative to evaluation (in this case being defined as systematic, pre-planned enquiry) which might be described as the "casual approach." The term "casual approach" is used to capture the essence of a typical mode or style that administrators, supervisors, and staff personnel use to understand events occurring about them. The intent here is not to be flippant and critical but rather examine more carefully a natural phenomenon of organizational life that affects all of us.

This practice which is being referred to as the "casual approach" appears to be a dominant style of a large percentage of organizational personnel. Its dimensions include face-to-face discussions, random observations, group discussions, and individual reflections. A breakdown of each of these factors might be described as follows:

- Face-to-face discussions

This approach includes chance or planned encounters where persons might discuss a wide-range of job and non-job related topics. It is a subtle process for the exchange of information, maintenance of work relationships and a potential resolver of organizational issues. It is the alternative to an evaluator's or researcher's use of an interview for gathering needed information. In part, qualitative evaluators advocate for such non-structured, open-ended interviews in order to better understand a person's experience within the organization rather than using a more formal, structured set of specific, close-ended questions. In any regard, a principal source for personnel (administrators and staff) are such encounters which are face-to-face but not limited to such since telephone conversations and memoranda do provide excellent supplements and often stimulate face-to-face contacts.

As with most approaches for developing insight, this approach has its advantages and disadvantages. First, some advantages include: it is a natural form of organizational communications blending with the customs and norms of the organization; it can be less threatening thus persons may be open to sharing true feelings and insights; and it can be two-way thus permitting an

exchange rather than a one-way lecture. Of course, each of these advantages can occur in the opposite direction. For example, some persons are better than others at face-to-face conversations and negative effects can results from such exchanges. In addition, such contacts can wander from the need or concern of the moment, become very superficial and be subject to subtle manipulation. Also, given the unplanned nature of this approach, it can result in chance encounters which create a bias perspective rather than provide a complete picture or balanced set of views. Thus, depending on the skills of the communicators, the climate of the organization for openness, and availability of encountering a broad spectrum of views, this strategy can produce a rather distorted set of insights.

• Random observation

Another tool widely used by personnel in organizations is random observations, a small portion of which may be preplanned but more likely being chance occurrences. In the field of education, unfortunately, conventional wisdom tends to place greater significance on first-hand observation than it possibly merits. Obviously, it has value in terms of being exposed to the mileu in which a person functions and being able to understand the realities in which ideals are compromised. However observers and observations do not always result in these understandings. Observers have their filters, may not try to obtain a balance of observations, may be distracted by unique and isolated factors, and may lack sufficient insights, experiences that allow the subtleties to be missed. There is no doubt the use of the powers of observation is an important source of information but simultaneously it can also be a major source for distorted information.

• Group discussions

Group discussion (two or more can constitute a group), albeit formal or informal, provides a forum for the exchange of information, opinions and ideas for addressing organizational concerns. Often the informal groups are beyond careful analysis since they occur out of view of most people in organizations and the substance of the discussions is seldom recorded. In any regard, something is known about them because they are part of most persons' experiences in an organization, the results of their discussions do surface particularly in formal group meetings, and direct glimpses are provided through coffee break and lunchroom discussions.

Formal group exchanges are generally more public, often their discussions are recorded, and frequently the results of these discussions are shared outside the group. However, it seems few people speak highly of having to attend formal group meetings (i.e., staff, task forces, committees, administrative committees, etc.) These meetings are often viewed as a waste of time, nothing very serious being accomplished at them, administrators do what they want regardless of input, meetings tend to be disorganized and so on. Fortunately or unfortunately, these forums do play a major role in the exchange of information and judgments. Again, depending on the quality of these meetings, how they

are organized, how persons are selected and given an opportunity to participate can greatly alter their outcomes and their potential value to organizational policy or decision-making.

• Individual reflections

Certainly out of sight of most persons in an organization is how individuals process their exposure to face-to-face encounters, observations, group meetings and written reports. It is truly Ashbey's black box. A general understanding can be surfaced as to the inputs and outputs of such reflections but the process is obscure unless some attempt is made by decision-makers to share as to how they arrived at their judgments. And, of course, such explanations do not guarantee an accurate representation of what exists or took place.

In sum, when a more careful examination is made of a principle alternative to formal systematic approaches to gathering insight about organizational and programmatic functioning, it raises some doubts about relying on casual approaches as a primary strategy for evaluating organization performance for decision-making, planning, and/or policy development.

3.0 Evaluation Activities

Some persons might define the preceding description as an evaluation activity and be willing to stop there. However, because of the serious limitations to the "casual approach," there is pressure to develop other alternative strategies. The casual approach will no doubt continue to be a major strategy for decision-makers and is appropriate as a sole approach where risks are low, scope of effort is limited, resources required are minimal, and the size of operation permits the easy "eyeballing" of the major activities. However, in conditions which call for greater accuracy and consistency of performance, resources required are high, operations are complex, and the social-political milieu demands greater accountability, these are conditions worthy of using more systematic, pre-planned strategies of evaluation. What is suggested here is that the potential strengths of the casual approach be strengthened with the tools and discipline of evaluation and research.

It may be of some assistance to briefly examine the distinctions between evaluation and research, to examine alternative strategies for evaluation, and four types of evaluation.

As is displayed in Exhibit 1, the strategies of evaluation, evaluation research and research can be portrayed as highly interrelated and the emphasis of each is primarily related to the developmental state of a program or project. This display tries to reduce the polarization often suggested among advocates of these major processes of enquiry or problem solving strategies that often fosters an "either-or" psychology. Also portrayed are specific types within each of the broad fields of enquiry which can occur at different points in time during the life of a program or project.

Exhibit 2 portrays in more detail some comparisons between these broad methodologies in regard to purpose, process, products, and criteria for judging quality. A fourth methodology (policy research) is added only for the purpose of providing a full spectrum of methodologies and to illustrate a new emerging field of enquiry. It is not central to the point of this paper at this juncture but does foreshadow a field that will be of greater interest in the future. As can be observed in Exhibit 2, there are distinct and somewhat different purposes for each methodology and these call for varying strategies, products and ultimate criteria for judging their payoff.

Not to confuse the topic further but rather to appreciate some subtle nuances within the field of evaluation itself, Exhibit 3 is provided. Depending on the major questions confronting an organization, a program or project, it is sufficient cause to vary the particular emphasis of the evaluation as is suggested by the models portrayed in Exhibit 3.

As these exhibits suggest, there is no one way for approaching an evaluation, there is no single set of activities which are limited to evaluation, and that the desired purposes for systematic enquiry must be part of the process of determining an appropriate set of activities.

4.0 Worth?

Determining if evaluation activities (assuming more formal approaches) are worth the effort and expenditure of resources is a hard question to answer. Again, it seems that all things tend to be relative in that some organizations have more or less of a need to conduct evaluations, have greater or lesser resources to allocate to evaluation activities, and have more or less clarity concerning standards of performance.

• Need

Need or desire to enter into formal evaluation activities can stem from numerous factors including the degree of external/internal pressure to perform such tasks, the degree to which the organization is threatened with competitors or image problems, the degree to which management of the organization value such systematic efforts, and the degree to which the organization faces open-ended problem situations.

Schools and public-supported higher education institutions are facing a much more skeptical and financially hard-pressed public who are seeking greater evidence of quality control procedures within these institutions. This phenomenon coupled with members of an organization having to compete for resources which are declining, create enormous pressure for a flurry of evaluation activities. A case example in a higher education organization was a particular residential subunit being flat out told to justify the added expenses required to run their unique programs. The central administration demand resulted in the design of an impact evaluation study.

Organizations which have competitors and/or organizations which have experienced some recent damage concerning their image are very likely organizations that will rely on evaluation activities, often involving external

consultants, in order to improve their status. This is not to suggest that evaluation activities can substitute for substance but it does suggest that part of the image is linked to an organization's capacity and willingness to solicit information about its performance.

Again, the need to conduct evaluations can be greatly increased by the complexities of issues and problems faced by the organization and the attitude of the managerial leadership of the organization toward the "value" of systematic problem-solving.

• Resources

The need to do more in the area of evaluation can be greatly compromised by an organization whose resources are significantly limited. It obviously requires time, energy, expertise, and some commitment to enter into an array of evaluation efforts. When an organization finds itself very busy putting out fires, staff being stretched beyond reason, and the presence of internal conflict, it is difficult to muster additional energy to address and follow through on a process of systematic enquiry. In other words, the need may be high for evaluation activities but the resources need to be expanded in order to permit some discretionary choices in this direction.

• Standards

There is no quicker way to create a flurry of evaluation activities than to create a set of fairly explicit standards for organization-wide or subunit performance. The greater amount of vagueness concerning performance indicators, the more difficult becomes the task of evaluation. Once a set of standards is articulated and once these standards have clout behind them (i.e., removal of accreditation, reduction in funds, and/or lost image), a significant increase in evaluation activities usually results.

5.0 A Further Note

A decision to embark on evaluation activities is linked to some mix of the need to do it, resources available, and the availability of explicit standards by which judgments can be rendered. There is, however, another part to the question of worth which is germain at this point, and that is a question concerning the state-of-the-art. Even if the three factors above were favorable for doing an evaluation, the question remains whether the technology of evaluation is sufficiently reliable to produce usuable, timely and accurate information at a reasonable cost. Some attempts have been made to spell out standards for the conducting of evaluation studies (see Exhibits 4 and 5) which can provide a framework for answering this question. It does seem that much has been learned over the past ten years which suggests that the state-of-the-art is more than a collection of concepts pirated from the field of research with little adaptation and sensitivity toward the unique features of applied settings. The theory and practice continue to grow and it seems real strides are being made in how to address the following areas:

1. Clarifying the audiences and purposes of an evaluation study.
2. Developing a partnership relationship between the evaluators and staff of programs and between evaluation and practice.
3. Adapting research methods including design, data gathering, data analysis which fit with the requirements of an evaluation study.
4. Developing reporting procedures for, during, and at the conclusion of a study which enable staff to understand and use such results.
5. Designing formats and procedures for program monitoring and documentation of program activities and results.

Areas which continue needing improved strategies are:

1. Providing adequate funds to support evaluation activities.*
2. Disseminating to a wider audience of people knowledge and skills associated with evaluation activities.
3. Integrating evaluation activities within the normal operations of an organization.
4. Integrating the technologies of qualitative and quantitative methodologies.
5. Reducing the skepticism of qualitative approaches to evaluation.
6. Determining more specifically how organizations and their members use knowledge gained from evaluation studies.**

6.0 Summary

In this paper has been presented a critique of the casual approach to evaluation which seems to be a common practice and which seems to be taken for granted as an effective tool. Also, a fuller explication has been provided in understanding the interface of evaluation and research and various approaches subsumed within evaluation. In addition, the dimensions which condition the worth of engaging in evaluation activities are presented. And, finally, a statement about the state-of-the-art in regard to the technology of conducting evaluation studies is presented.

Hopefully this commentary provides some challenge to the status quo and provides impetus to learn more about the field of evaluation or at least to appreciate circumstances which may merit evaluation efforts. At least minimally those who will continue to rely on the casual approach will become better

---

* Somehow the investment in evaluation activities needs to be viewed as cost saving and not an additional expense. It seems highly probable that evaluation activities can contribute to eliminating or modifying costly and ineffective procedures.

**CEPR presently follows a practice of gaining feedback on products produced for its clients in order to determine more carefully how such efforts have payed off.

informed as to how to make their observations and face-to-face contacts more than chance occurrences with little forethought as to what to look for or what to discuss. Certainly there are times for spontaneity of action but hopefully these can be balanced by more carefully planned methods of determining organization and individual performances. For it is very clear, evaluation activities by their very nature do force a review of values and practices in an organization and their interrelationship.

Related Readings

Alkin, Marvin C., Daillak, Richard, and White, Peter. Using Evaluations. Beverly Hills, CA: Sage Publications, 1979.

Cook, Thomas D., and Reichardt, Charles S. Qualitative and Quantitative Methods in Evaluation Research. Beverly Hills, CA: Sage Publications, 1979.

Dolbeare, Kenneth, editor. Public Policy Evaluation. Beverly Hills, CA: Sage Publications, 1975.

House, Ernest R. Evaluating with Validity. Beverly Hills, CA: Sage Publications, 1980.

Morris, Lynn Lyons and Fitz-Gibbon, Carol. Evaluator's Handbook. Beverly Hills, CA: Sage Publications, 1978.

Patton, Michael Q. Utilization-Focused Evaluation. Beverly Hills, CA: Sage Publications, 1978.

Popham, W. James, editor. Evaluation in Education. Berkeley, CA: McCutchan Publishers, 1974.

Rossi, Peter H., Freeman, H.E., Wright, S.R. Evaluation: A Systematic Approach. Beverly Hills, CA: Sage Publications, 1979.

Rutman, Leonard. Planning Useful Evaluations. Beverly Hills, CA: Sage Publications, 1980.

Journals of Interest:

CEDR Quarterly (Center on Evaluation, Development and Research); Phi Delta Kappa

Educational Evaluation and Policy Analysis; American Educational Research Association

Evaluation and Program Planning; Pergamon Press

Evaluation News; Sage Publications

Evaluation Review; Sage Publications

Knowledge: Creation, Diffusion, Utilization; Sage Publications

Policy Studies Journal; Policy Studies Organization, University of Illinoia

Exhibit 1

Interrelationship of Evaluation, Evaluation Research and Research to Program/Project Development

Interrelationship of Evaluation, Evaluation Research and Research to Program/Project Development

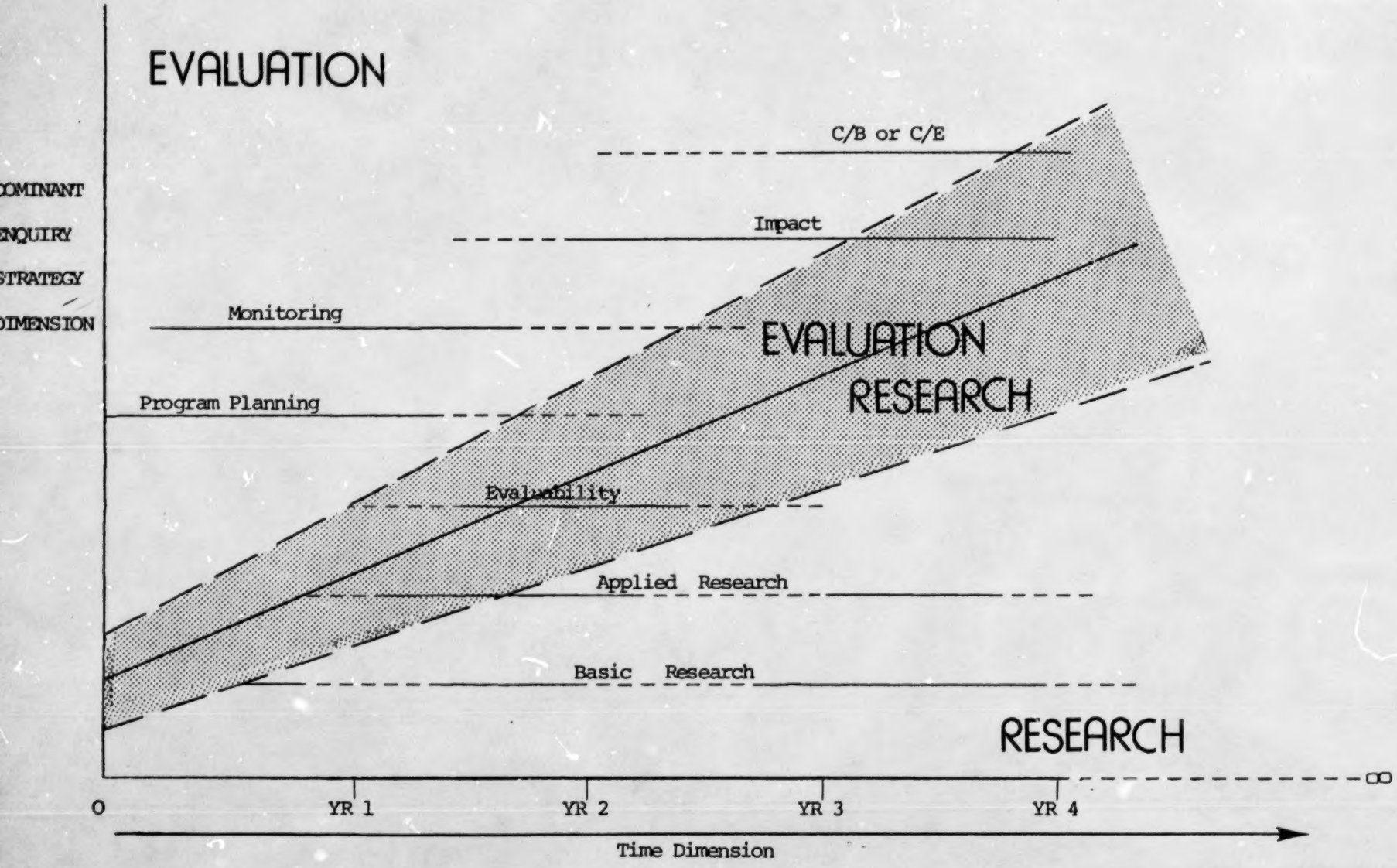

Exhibit 2

Comparison of Alternative Problem Solving Methodologies

Comparison of Alternative Problem Solving Methodologies

	RESEARCH	EVALUATION	EVALUATION RESEARCH	POLICY ANALYSIS/ RESEARCH
URPOSE	build society's knowledge base	facilitate program decision making and improvement	prove the worth of some social activity	aid in development of institutional policies
ROCESS	problem identification	decision and/or need identification	determine purposes evaluability assessment -describe program -analyze program description -construct alternatives -implement findings	Delphi studies
	research design	decision and program parameter specification		
	data generation	evaluation design (options) -program planning -program monitoring -impact analysis -C/B or C/E analysis	specification of variables -identify program components -specify goals and/or effects -specify antecendent variables -specify intervening variables	experimental & quasi-experimental designs scenario development forecasting studies judicial proceedings cost/effectiveness studies
	data analysis	data & information generation	development of measures	
	reporting	data & information analysis interpretation & reporting	data analysis interpretation & reporting	
RODUCTS	generalizeable knowledge	information for a specific decision and/or program	information for determining merit, quasi-generalizeable knowledge	identification of social choices constraints clarified
RITERIA	standards of scientific excellence Theory based	decision relevance, comprehensiveness, credibility, timeliness, efficiency, and standards of scientific objectivity	relevance, comprehensiveness, credibility, timeliness, efficiency and internal/ external validity	relevance, comprehensiveness, credibility, timeliness, political viability, efficiency, and scientific objectivity

Exhibit 3

A Taxonomy of Major Evaluation Approaches

A TAXONOMY OF MAJOR EVALUATION APPROACHES *

Model	Major Audiences or Reference Groups	Assumes Consensus on	Methodology	Outcome	Typical Questions
Systems Analysis	Economists, managers	Goals, known cause and effect, quantified variables	PPBS, linear programming, planned variation, cost-benefit analysis	Efficiency	Are the expected effects achieved? Can the effects be achieved more economically? What are the most efficient programs?
Behavioral Objectives	Managers, psychologists	Prespecified objectives, quantified outcome variables	Behavioral objectives, achievement tests	Productivity, accountability	Is the program achieving the objectives? Is the program producing?
Decision-making	Decision-makers, especially administrators	General goals, criteria	Surveys, questionnaires, interviews, natural variation	Effectiveness, quality control	Is the program effective? What parts are effective?
Goal-free	Consumers	Consequences, criteria	Bias control, logical analysis, modus operandi	Consumer choice, social utility	What are all the effects?
Art Criticism	Connoisseurs, consumers	Critics, standards	Critical review	Improved standards, heightened awareness	Would a critic approve this program? Is the audience's appreciation increased?
Professional Review	Professionals, public	Criteria, panel, procedures	Review by panel, self study	Professional acceptance	How would professionals rate this program?
Quasi-legal	Jury	Procedures & Judges	Quasi-legal procedures	Resolution	What are the arguments for and against the program?
Case Study	Client, practitioners	Negotiations, activities	Case studies, interviews, observations	Understanding diversity	What does the program look like to different people?

*from Ernest House, Evaluating With Validity. Beverly Hills: Sage Publications, 1980, p. 23.

Exhibit 4

Standards for Evaluations of Educational Programs

UTILITY STANDARDS*

SUMMARY OF THE STANDARDS

A Utility Standards

The Utility Standards are intended to ensure that an evaluation will serve the practical information needs of given audiences. These standards are:

A1 Audience Identification

Audiences involved in or affected by the evaluation should be identified, so that their needs can be addressed.

A2 Evaluator Credibility

The persons conducting the evaluation should be both trustworthy and competent to perform the evaluation, so that their findings achieve maximum credibility and acceptance.

A3 Information Scope and Selection

Information collected should be of such scope and selected in such ways as to address pertinent questions about the object of the evaluation and be responsive to the needs and interests of specified audiences.

A4 Valuational Interpretation

The perspectives, procedures, and rationale used to interpret the findings should be carefully described, so that the bases for value judgments are clear.

A5 Report Clarity

The evaluation report should describe the object being evaluated and its context, and the purposes, procedures, and findings of the evaluation, so that the audiences will readily understand what was done, why it was done, what information was obtained, what conclusions were drawn, and what recommendations were made.

A6 Report Dissemination

Evaluation findings should be disseminated to clients and other right-to-know audiences, so that they can assess and use the findings.

A7 Report Timeliness

Release of reports should be timely, so that audiences can best use the reported information.

A8 Evaluation Impact

Evaluations should be planned and conducted in ways that encourage follow-through by members of the audiences.

*from "Standards for Evaluations of Educational Programs, Projects, and Materials," The Joint Committee on Standards for Educational Evaluation, New York: McGraw-Hill Company, 1981.

B

FEASIBILITY STANDARDS

SUMMARY OF THE STANDARDS

B Feasibility Standards

The Feasibility Standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal; they are:

B1 Practical Procedures

The evaluation procedures should be practical, so that disruption is kept to a minimum, and that needed information can be obtained.

B2 Political Viability

The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted.

B3 Cost Effectiveness

The evaluation should produce information of sufficient value to justify the resources expended.

C PROPRIETY STANDARDS

SUMMARY OF THE STANDARDS

C Propriety Standards

The Propriety Standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results. These standards are:

C1 Formal Obligation

Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that these parties are obligated to adhere to all conditions of the agreement or formally to renegotiate it.

C2 Conflict of Interest

Conflict of interest, frequently unavoidable, should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results.

C3 Full and Frank Disclosure

Oral and written evaluation reports should be open, direct, and honest in their disclosure of pertinent findings, including the limitations of the evaluation.

C4 Public's Right to Know

The formal parties to an evaluation should respect and assure the public's right to know, within the limits of other related principles and statutes, such as those dealing with public safety and the right to privacy.

C5 Rights of Human Subjects

Evaluations should be designed and conducted, so that the rights and welfare of the human subjects are respected and protected.

C6 Human Interactions

Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation.

C7 Balanced Reporting

The evaluation should be complete and fair in its presentation of strengths and weaknesses of the object under investigation, so that strengths can be built upon and problem areas addressed.

C8 Fiscal Responsibility

The evaluator's allocation and expenditure of resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible.

D

ACCURACY STANDARDS

SUMMARY OF THE STANDARDS

D Accuracy Standards

The Accuracy Standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features of the object being studied that determine its worth or merit. These standards are:

D1 Object Identification

The object of the evaluation (program, project, material) should be sufficiently examined, so that the form(s) of the object being considered in the evaluation can be clearly identified.

D2 Context Analysis

The context in which the program, project, or material exists should be examined in enough detail, so that its likely influences on the object can be identified.

D3 Described Purposes and Procedures

The purposes and procedures of the evaluation should be monitored and described in enough detail, so that they can be identified and assessed.

D4 Defensible Information Sources

The sources of information should be described in enough detail, so that the adequacy of the information can be assessed.

D5 Valid Measurement

The information-gathering instruments and procedures should be chosen or developed and then implemented in ways that will assure that the interpretation arrived at is valid for the given use.

D6 Reliable Measurement

The information-gathering instruments and procedures should be chosen or developed and then implemented in ways that will assure that the information obtained is sufficiently reliable for the intended use.

D7 Systematic Data Control

The data collected, processed, and reported in an evaluation should be reviewed and corrected, so that the results of the evaluation will not be flawed.

D8 Analysis of Quantitative Information

Quantitative information in an evaluation should be appropriately and systematically analyzed to ensure supportable interpretations.

D9 Analysis of Qualitative Information

Qualitative information in an evaluation should be appropriately and systematically analyzed to ensure supportable interpretations.

D10 Justified Conclusions

The conclusions reached in an evaluation should be explicitly justified, so that the audiences can assess them.

D11 Objective Reporting

The evaluation procedures should provide safeguards to protect the evaluation findings and reports against distortion by the personal feelings and biases of any party to the evaluation.

Exhibit 5

Standards for Program Evaluation

FORMULATION AND NEGOTIATION *

Before an evaluation project or program is undertaken, the concerned parties should have a clear mutual understanding of what is to be done, how it is to be done, and why, and an appreciation of possible constraints or impediments. The knowledge available will vary widely and affect the manner of compliance with the standards in this section, but a significant effort and resource commitment will usually provide an important framework for an effort that is technically sound, relevant, and justifiable in terms of costs.

1. The purposes and characteristics of the program or activity to be addressed in the evaluation effort should be specified as precisely as possible.

2. The clients, relevant decisionmakers, and potential users of the evaluation results should be identified, and their information needs and expectations made clear. Where appropriate, evaluators should also help identify areas of public interest in the program.

3. The type of evaluation effort required should be identified and its objectives made clear; the range of activities to be undertaken should be specified. (See categories A - F, Introduction)

4. An estimate of the cost of the proposed evaluation effort and, where appropriate, of alternatives should be provided; this estimate should be prudent, ethically responsible, and based on sound accounting principles.

5. Agreement should be reached at the outset that the evaluation is likely to produce information of sufficient information value, applicability, and potential for utilization to justify the resources used.

*from "Standards for Program Evaluation, Exposure Draft," Evaluation Research Society, May 1980.

6. The feasibility of undertaking the evaluation effort should be estimated either informally or through formal evaluability assessment (see p. 5). (Some of the factors to consider are the clarity of the program description and objectives; prospects for needed cooperation; the plausibility of postulated cause-effect relationships; the availability of time, money, and expertise to carry out the evaluation; and administrative, fiscal, and legal constraints).

7. Restrictions, if any, on access to the data and results from an evaluation should be clearly established and agreed to between the evaluator and the client at the outset. (In some cases; e.g., government-sponsored studies where freedom of information statutes apply, and where it is understood that the results of evaluation studies automatically go into the public domain, the right-to-know question is not negotiable. The sponsor or evaluator is obligated to point this out at the beginning. In other cases; e.g., confidential studies undertaken for private individuals and organizations, the client may rightfully expect the confidentiality of the findings to be maintained.)

8. Potential conflicts of interest should be identified and steps should be taken to avoid compromising the evaluation processes and results.

9. Respect for and protection of the rights and welfare of all parties to the evaluation should be a central consideration in the negotiation process.

10. Accountability for the technical and financial management of the evaluation once it is undertaken should be clearly defined.

11. All agreements reached in the negotiation phase should be specified in writing, including: schedule; obligations and involvements of all formal parties to the evaluation and of all parties with specific roles to play in the effort; policies and procedures on access to the data. Provision should be made for subsequent agreements and these should be specified in writing.

12. Evaluators should not accept obligations that exceed their professional qualifications or the resources available to them.

STRUCTURE AND DESIGN

The design for any evaluation cannot be conceived in a vacuum. It is necessarily influenced by logistical, ethical, political, and fiscal concerns, and therefore must take these into account as well as methodological requirements. This applies to each type of evaluation specified in the introductory section. Some of the principal concerns that extend beyond methodological requirements and influence the design itself are embodied in standards 1, 2, 3, 6, 7, 8, and 9. Designs will vary in rigor and not all instruments are equally objective. However, even with these broad variations, the following standards generally apply. For example, the approach to a case study is as subject to specification as the design of an experimental study; the reliability of judgments is as much at issue as the reliability of objective tests.

13. For all types of evaluations, a clear approach or design should be specified and justified as appropriate to the types of conclusions and inferences to be drawn.

14. For impact studies, the central evaluation design problem of estimating the effects of non-treatment, and the choice of a particular method for accomplishing this, should be fully described and justified.

15. If sampling is to be used, the details of the sampling methodology (choice of unit, method of selection, time frame, etc.) should be described and justified, based on explicit analysis of requirements of the evaluation, including generalization beyond the population sampled.

16. The measurement methods and instruments should be specified and described, and their reliability and validity of application to the characteristics to be measured should be estimated.

17. Justification should be provided that the best and most appropriate procedures and instruments have been specified.

18. The necessary cooperation of program staff, affected institutions and members of the community, as well as those directly involved in the evaluation, should be planned, and assurances obtained should be noted (see standard 11).

DATA COLLECTION AND PREPARATION*

These standards assume that data collection is carried out within the specifications of a sound design and plan of work. (see standards 1-18) However, at the time the data collection methods are specified it is usually not too late to make some reasonable changes in the design that accomodate to unforeseen circumstances. If insurmountable logistical difficulties or other adverse circumstances arise, the design and work plan should be revised accordingly.

19. A data collection and preparation plan should be developed in advance of data collection.

20. Provision should be made for the detection, reconciliation, and documentation of departures from the original design.

21. Evaluation staff should be selected, trained, and supervised to ensure competence, consistency, impartiality, and proper ethical practice.

22. All data collection activities should be conducted so that the rights, welfare, dignity, and worth of individuals are respected and protected.

23. The estimated validity and reliability of data collection instruments and procedures (see standard 16) should be verified under the prevailing circumstances of their use.

24. Analysis of the sources of error should be undertaken, and provisions for quality assurance and control should be established to adequately meet the requirements of the overall design and anticipated data analyses.

---

*Where secondary data are used, the evaluator can describe what is known about whether these standards have been met by the processes through which the data were originally produced.

25. The data collection and preparation procedures should provide safeguards so that the findings and reports are not distorted by any biases of data collectors.

26. Data collection activities should be conducted with minimum disruption to the program under study and with minimum imposition on the organizations or persons from whom data are gathered.

27. Procedures that may entail adverse effects or risks should be subjected to external independant review and then used only with informed consent of the parties affected.

28. Data should be handled and stored so that release to unauthorized persons is prevented, and access to individually identifying data is limited to those with a need to know. (see standard 7)

29. Documentation should be maintained of the source, method of collection, circumstances of collection, and processes of preparation for each item of data.

30. Appropriate safeguards should be employed to insure against irrecoverable loss of evaluation data through catastrophic events.

DATA ANALYSIS AND INTERPRETATION

The choice of analytic procedures, like the choice of data collection methods, is more or less dictated by the structure and design of the evaluation effort. At the data analysis stage, the evaluator no longer has freedom of action to change the design and has the requirement to temper the analyses to peculiarities of the data actually collected. New methods and procedures, some detailed and rigorous, are appearing in the literature and evaluators may want to include some of these among methods that are specified and used.

31. The analytic procedures should be matched to the general purposes of the evaluation, the design, and the data collection.

32. All analytic procedures, along with their underlying assumptions and limitations, should be described explicitly, and the reasons for choosing the procedures should be clearly explained. (The level of detail required in the descriptions will vary with the familiarity of the procedure to the primary audience; the better known the procedure, the shorter the description needed.)

33. Analytic procedures should be appropriate to the properties of the measures used and to the quality and quantity of the available data.

34. The units of analysis should be appropriate to the way the data were collected and the types of conclusions to be drawn.

35. Justification should be provided that the best and most appropriate analytic procedures have been applied.

36. Documentation should be adequate to make the analyses replicable.

37. When quantitative comparisons are made (e.g., x is greater than y) tests of statistical significance should be applied and interpretations should be stated with some indication of confidence.

38. Cause-and-effect interpretations should be bolstered not only by reference to the design but also by recognition and elimination of plausible rival explanations.

39. Findings should be reported in a manner that distinguishes among objective findings, opinions, judgments, and speculation.

COMMUNICATION AND DISCLOSURE

Good communication is obviously essential to a well-formulated and executed evaluation effort and to any utilization of the results. In particular, good communication is necessary to clarify the nature of the program, the expectations for the evaluation, and even the type of evaluation effort required (see standards 1, 2, and 3); to anticipate restrictions on release of results and potential conflicts of interest (see standards 7 and 8); to establish accountability for the effort (see standards 10 and 11); to secure the cooperation of parties involved in the program and evaluation (see standards 18 and 27); and to distinguish clearly objective findings and other information (see standard 39). Particularly in the teaching role a large portion of the evaluator's time is devoted to communication. In short, communication is not to be equated solely with the "final report." However, most evaluation efforts do produce certain formal reports, intermediate and final, written and oral, and there are standards these reports should meet.

40. Findings should be presented clearly, completely, and fairly.
(see standard 39.)

41. Findings should be organized and stated in language understandable by decisionmakers and other audiences and any recommendations should be clearly related to the findings.

42. Findings and recommendations should be presented in a framework that indicates their relative importance.

43. Assumptions should be explicitly acknowledged.

44. Limitations caused by constraints on time, resources, data availability, etc. should be stated, (see standards 5, 6, 7, 11, and 12). (Suggestions should be included on how to study those issues and questions that need further study and encouragement or assistance in doing so should be offered).

45. Complete explanation and description of how findings and results were derived should be accessible.

46. Persons, groups, and organizations who have contributed to the evaluation effort should receive feedback appropriate to their indicated perspectives and applications.

47. Disclosure should follow the legal and proprietary understandings agreed upon in advance (standard 7), with the evaluator serving as a proponent for the fullest, most open disclosure appropriate.

48. Officials authorized to release the evaluation data should be specified.

49. The finished data base and associated documentation should be organized in a manner consistent with the accessibility policies and procedures. (See standards 7, 28, 29, 32, and 36).

UTILIZATION

The usual reason for conducting an evaluation is a functional one: to help those affected to be better informed about the feasibility of undertaking the program, the reasonableness of evaluating it, the program operation and its effects, and the results of previous evaluation efforts. Utilization cannot be guaranteed, of course, but it will be more likely if careful attention is given to the information needs of the potential users of the results throughout all phases of the evaluation (see especially standards 2, 3, 18, 40-46). Beyond the day-to-day processes of encouraging responsiveness to the evaluation, there are some other considerations that the evaluator needs to keep in mind.

50. Evaluation results should be made available to appropriate users before relevant decisions must be made.

51. Evaluators should try to anticipate and prevent misinterpretations and misuses of evaluative information.
(The evaluator, of course, cannot be held responsible for misuses of evaluative information. Nevertheless, follow-up contacts with users, rebuttals of misinterpretation, and promotion of an open exchange of information should be a part of the evaluator's responsibility.)

52. The evaluator should bring to the attention of decisionmakers and other relevant audiences suspected side effects--positive or negative--of the evaluation process.

53. Evaluators should clearly distinguish between the findings of the evaluation and any policy recommendations based on them.
(If evaluators are called upon to go beyond the findings and to make policy recommendations, or if they initiate such recommendations, they must make clear the difference between recommendations that are clearly supported by the findings of the evaluation and other recommendations and support for them.)

54. In making recommendations about corrective courses of action evaluators should carefully consider and indicate what is known as a basis for estimating the probable effectiveness and costs of the recommended courses of action.

55. Evaluators should maintain a clear distinction between their role as an evaluator and any advocacy role they choose to adopt.
(Evaluators should be aware of the apparent conflict between advocating certain positions and presenting evaluation results. Evaluators may wish to take advocacy stands, but when they do they should not assume that they possess any special status or competence.)